

Retrospective Conversion of Masters Theses at Brigham Young University: Two Models

Scott Eldredge - Brigham Young University

May 30, 2002

The Projects

- *Mormon Theses*
- *Benson Institute Theses*

Mormon Theses

Selection

- *Limited by topic*
- *Importance to institution*
- *User demand*

Rights Management

- *Permission forms*

Outsourcing

- ***Vendor Selection***
 - RFI
 - RFP
 - **Proposal Review**

Brigham Young University
Harold B. Lee Library

**REQUEST FOR INFORMATION (RFI)
FOR DIGITAL IMAGING PRODUCTION SERVICES**

March 7, 2000

The purpose of this Request for Information is to enable Brigham Young University Library to identify vendors who would like to receive a Request for Proposal (RFP) to bid on a digital conversion project involving the creation of 150,000 digital images and associated metadata. BYU will digitize a collection of thesis materials relating to Mormonism. Up to 700 volumes (150,000 pages) will be selected and scanned in a 10 month period (March 2000 - January 2001). BYU is seeking a vendor (or group of vendors) to perform the digital conversion, metadata creation (directory structuring, file naming, header information input), text conversion (OCR/re-keying), SGML markup, and to produce printouts of the digital files.

Please note the following important dates associated with the RFI and RFP:

March 24, 2000	Vendor responses to RFI due.
April 28, 2000	Brigham Young University Library sends the Request for Proposal (RFP) to vendors.
May 12, 2000	Vendors notify Brigham Young University Library indicating whether they will be responding to the RFP and wish to receive the sample volumes for scanning.
May 22, 2000	Brigham Young University Library supplies each vendor with sample volumes and accompanying instructions for the Preliminary Production Test.
June 27, 2000	Vendors deliver to Brigham Young University their RFP response, plus the sample volumes and the products of the preliminary test.
Week of July 10, 2000	RFP responses evaluated and Vendor(s) chosen.
Week of July 17, 2000	Contract(s) awarded, and unsuccessful vendors notified.
July 28, 2000	Shipment #1 consisting of 5,000 pages sent from BYU to vendor(s).
August 18, 2000	Shipment #1 products returned from vendor(s) to BYU.

Brigham Young University
Harold B. Lee Library

**REQUEST FOR INFORMATION (RFI)
FOR DIGITAL IMAGING PRODUCTION SERVICES**

RESPONSE FORM

1. During this 9-month project (March 2000 – December 2000), nearly 150,000 pages will be processed, including scanning, metadata creation, text conversion, SGML encoding, and printing. Therefore, we anticipate shipping 25,000 pages/month to vendor(s). Please confirm your ability to meet such a production workload.
2. We anticipate that the majority of images can be captured as 600 dpi bitonal images. Some of the volumes contain halftones that will require grayscale scanning or enhancement if captured bitonally. Please indicate your capability to scan images as:

_____ 600 dpi bitonal, with enhancement for halftone capture (please describe)

_____ 300 dpi, 8-bit gray

_____ 400 dpi, 8-bit gray

3. BYU may choose not to disbind the volumes prior to shipment to the vendor(s).* In this case, the majority of them will have to be scanned from bound volumes. We anticipate that the pages of any disbound volumes will be brittle and must be handled carefully - placed one page at a time on the scanning platen and that automatic document feed handlers will not be used for most of the volumes. Please indicate whether you have experience doing the following:

_____ Manually scanning one page at a time on a flatbed scanner

_____ Scanning bound volumes

* The full RFP will request separate bids for scanning from bound volumes and from unbound or photocopied sheets. The bid figures will influence BYU's decision regarding disbinding.

4. The physical dimensions of pages will vary from 8.5" x 11" to 17" x 22". Please indicate the page dimensions you can handle:

____ 8.5"x 11" ____ 8.5" x 14" ____ 11"x 17" ____ 17" x 22"

5. Bitonal and grayscale images should be captured as TIFF 6.0 files with Intel byte order, and compressed using ITU Group 4 (formerly CCITT). Please confirm that you can deliver images adhering to these file formats and compression techniques.

6. BYU will require the vendor(s) to create associated metadata to facilitate the matching of image files (file names) with physical volumes, and to provide bibliographic access to the files in several levels. Metadata creation will include directory structuring, file naming, indexing, and TIFF header data input. Please indicate your capability to provide these services.

7. All the images created during this project will be converted to text (ASCII) through OCR processing or rekeying at 99% accuracy level (maximum 1 error in every 100 words). The ultimate goal is to encode the textual information for indexing and other text manipulation purposes using the SGML TEI Lite DTD. Please indicate your capability to provide OCR/rekeying and SGML encoding services.

Outsourcing

- ***Vendor Selection***
 - RFI
 - RFP
 - **Proposal Review**

REQUEST FOR PROPOSAL (RFP) FOR DIGITAL IMAGING PRODUCTION SERVICES

Please note the following important dates associated with the RFP:

Friday, May 12, 2000, 5PM

Notify Brigham Young University Library indicating whether you will be responding to this RFP and wish to receive the sample volumes for scanning.

Monday, May 22, 2000

Brigham Young University Library will supply each vendor with sample volumes and accompanying instructions for the Preliminary Production Test.

Tuesday, June 27, 2000, 5PM

Vendors must deliver to BYU University their RFP response, plus the sample volumes and the products of the preliminary test.

Week of July 10, 2000

RFP responses evaluated and Vendor(s) chosen.

Week of July 17, 2000

Contract(s) awarded, and unsuccessful vendors notified.

Friday, July 28, 2000

Shipment #1 sent to vendor(s).

BRIGHAM YOUNG UNIVERSITY LIBRARY
REQUEST FOR PROPOSAL (RFP)
FOR DIGITAL IMAGING PRODUCTION SERVICES

I. INTRODUCTION

Brigham Young University (BYU) will undertake a project to digitize master's theses related to Mormonism. Up to 700 volumes (representing 150,000 images) will be selected, scanned, converted to text, SGML encoded, and made available online via a web-based interface. BYU is seeking a vendor or group of vendors to perform the digital conversion, metadata creation (directory structuring, file naming, header information, indexing), text conversion (OCR/rekeying), and SGML encoding/tagging. BYU will create access derivatives from the master images, therefore this proposal does not include any guidelines for the creation of derivatives.

The purpose of this Request for Proposal (RFP) is to enable BYU to identify and select a vendor (or vendors) to produce high quality bitmapped images of theses. During a 10 month period (March 2000 – January 2000), BYU will identify, prepare, disbind or photocopy (when necessary), and ship up to 700 volumes to the vendor(s) for the production of digital images. The vendor(s) will scan all pages as bitonal (one-bit) images with no gray scale or color scanning at a resolution of true 600 dpi. Halftone illustrations and finely detailed line art should be captured so as to suppress aliasing distortions and moiré patterns. Images will be produced in TIFF 6.0 and compressed using ITU Group 4 (formerly CCITT), Intel byte order.

The vendor will also produce associated metadata. In addition, the vendor will convert the digital images to text files via Optical Character Recognition (OCR) process or rekeying as necessary, and encode/tag the resulting text files using Standard Generalized Markup Language (SGML). BYU also requires the vendor to produce 600 dpi paper prints of the digital files on acid free paper to facilitate quality inspection of the scanned images and to create hard copy replacements for any deteriorating originals.

II. PROJECT OVERVIEW

BYU offers strong computing and networking resources and has made substantial contributions to the development of digital library initiatives, including participation in multi-institutional projects such as the American Memory Project. In this project BYU is embarking on its first large-scale effort in conversion and access of full text materials. BYU has experience with in-house scanning and is strongly supportive of the development of high quality, third party conversion services that can meet the needs of research libraries.

III. PROPOSAL SUBMISSION INFORMATION

A. Requirements for Submission

1. Submitted proposals must be addressed to:
Brigham Young University Purchasing Department
Attn: Michael Bitter
Purchasing Agent
C-173
Provo, UT 84602
(Phone: 801/378-3471 Fax: 801/378-7455)

A. Receiving Material from BYU

The vendor(s) must acknowledge receipt of each item in a shipment, using an annotated copy of the packing slip. If any discrepancies to the packing slip are found, or the originals and/or accompanying metadata instructions are problematic, the vendor needs to inform the institution immediately. All materials will be stored in a secure, dry location at the vendor(s), and great care should be taken in handling the fragile originals.

B. Scanning Procedures

1. System calibration and performance

Vendor shall exercise rigorous quality control to maintain consistency of output as described in ANSI/AIIM MS44-1988 (R1993), *Recommended Practice for Quality Control of Image Scanners*. Vendor shall ensure that the scanning system is free of dust and other distorting particles, that it maintains calibration throughout each shift, and that the appropriate technical targets are used. These targets shall include, but are not necessarily limited to:

RIT Alphanumeric Test Object

The 3 X 3 (1.9, 1.3, .7 Density Target) (Catalog No. A200-610-D1.9) may be ordered from:

T&E Center
Rochester Institute of Technology
Lomb Memorial Drive
Rochester, NY 14623-5604
(716) 475-2739

(Note: MS44-1988 does not mention the RIT Alphanumeric Test Object.)

AIIM Scanner Test Chart #2

The target (Catalog No. X4410) may be ordered from:

AIIM Bookstore
1100 Wayne Avenue, Suite 1100
Silver Spring, MD 20910-5603
(301) 587-8202

2. Capability to manage textual data for multipage volumes

Bound volumes for scanning consist of commercially bound theses. Most theses will be contained in single volumes. In a few cases, the a single thesis will be bound in two separate physical volumes

3. Setting up the theses

Prior to scanning, each thesis should be reviewed to determine the presence and nature of illustrations, the page dimensions, and the physical condition. The inspection will ensure that the scanner settings selected (threshold, filters, screens, TRC maps, page trim) will provide the best possible image capture for a given volume. These settings should be recorded on the accompanying worksheet.

Outsourcing

- ***Vendor Selection***
 - RFI
 - RFP
 - **Proposal Review**

Mormon Theses Proposal Evaluation Form

DC Labs		Score
Understanding of and compliance with the requirements of the RFP		1 2 3 4 5
Excellence of response		1 2 3 4 5
Ability to meet Preliminary Digital Production Test requirements		1 2 3 4 5
Satisfying mandatory technical requirements		1 2 3 4 5
Ability to meet the required scanning, metadata creation, text conversion, SGML encoding, and printout production within the required project timetable		1 2 3 4 5
Fair pricing of the proposal relative to other proposals received		1 2 3 4 5
Qualifications and experience, evidenced by customer references, resumes of key personnel, etc		1 2 3 4 5
Guarantee of work, and the nature and extent of vendor support		1 2 3 4 5
Financial stability and other various business issues as outlined in the RFP		1 2 3 4 5
Total Score		

Document Prep

- *Pulling*
- *Checkout*
- *Review – PNT*
- *Disbinding*
- *Packaging*
- *Shipping*
- *Receiving*
- *Review*

Production Note

# of pages to be scanned <small>includes PNT</small>	155
Root Identifier	0000lars
Number of Volumes	1

Scanner Settings	
threshold	
filters	
screens	
TRC maps	
page trim	

Bibliographic Information	
title control number	DGR1142
LC number (lc)	HQ3.02.L3771967
	Factors in the acceptance and adoption of family home evening in the L.D.S. church : a study of planned change
thesis title (mt)	
other thesis title (ot)	
thesis author (au)	Larson, Robert Ernest
secondary thesis author (sa)	
date of submission (yr)	1967
LC subject headings (su)	Mormon church--Family home evening program.
additional information (ad)	

Document Structures

page number range	label	image number range
000a	PNT	
r001, r002	TPG	
r01a	BLP	
r003	ACK	
r004-r008	TOC	
r009-r010	LOT	
0104-0110	BIB	
0139-0142	ABS	

Potential Grayscale Images

page number	type

Anomalies/Special Instructions

Note: The page right after 0110 is 110a.



Document Prep

- *Pulling*
- *Checkout*
- *Review – PNT*
- *Disbinding*
- *Packaging*
- *Shipping*
- *Receiving*
- *Review*

Document Deprep

- *Reprinting (Optional)*
- *Rebinding*
- *Re-tag (bar-code, etc.)*
- *Check in*
- *Reshelving*

Billing

- *Work within the system*
- *Review and approve invoices*
- *Track Budgets*

Conversion

- *Receipt of files*
 - FTP
 - CD

Quality control

- *Image files*
- *Text files*
- *Mark up*

Access

- *Decisions driven by capabilities*
- *Hyperion*
- *Convert to HTML*
 - XSL style sheet
- *JPEG surrogate files*
- *OPAC records*
 - 856 field
 - Separate record?

Benson Institute

ELECTRONIC THESES
2002
AND DISSERTATIONS

Selection

- *Decisions made by Institute*
- *Focus on developing countries*

Rights Management

- *Secured at time of submission*
- *International copyright laws apply*

Document Prep

- *Pulling*
- *Transcription (abstract)*
- *Translations (abstract)*
- *Review*
- *Disbinding*

Scanning

- *In house*

Quality Control

- *Post-scan review*

Document Deprep

- *Rebinding*
- *Return to Institute*
- *No copies in library*

Post Scan processing

- *Despeckle/Deskew*
- *Create PDFs*

Metadata

- *MARC record in OPAC*
- *URL in MARC record (856 field)*
- *Metadata for CONTENTdm from MARC record*

Load into Digital Library

- ***CONTENTdm***

Major Issues

- Rights Management
- In house vs. outsource
- Standards
 - Mark-up
 - Scanning
 - Metadata
- Access
- Digital Library Products or Databases
- Bandwidth considerations
- Costs

URLS

- ***Mormon Theses***

- <http://www.lib.byu.edu/hbli/digcol.htm>

- ***Benson Institute***

- <http://www.lib.byu.edu/hbli/benson/search.htm>